Nucleic Acids Research, 1995, Vol. 23, No. 21 4383–4390

Isolation and sequence analysis of *rpoH* genes encoding σ^{32} homologs from gram negative bacteria: conserved mRNA and protein segments for heat shock regulation

Kenji Nakahigashi, Hideki Yanagi and Takashi Yura*

HSP Research Institute, Kyoto Research Park, Kyoto 600, Japan

Received July 19, 1995; Revised and Accepted September 21, 1995

DDBJ accession nos D50828–D50832 (incl.)

ABSTRACT

The *rpoH* genes encoding homologs of *Escherichia coli* σ^{32} (heat shock σ factor) were isolated and sequenced from five gram negative proteobacteria (γ or α subgroup): *Enterobacter cloacae* (γ), *Serratia marcescens* (γ), *Proteus mirabilis* (γ), *Agrobacterium tumefaciens* (α) and *Zymomonas mobilis* (α). Comparison of these and three known genes from *E.coli* (γ), *Citrobacter freundii* (γ) and *Pseudomonas aeruginosa* (γ) revealed marked similarities that should reflect conserved function and regulation of σ^{32} in the heat shock response. Both the sequence complementary to part of 16S rRNA (the 'downstream box') and a predicted mRNA secondary structure similar to those involved in translational control of σ^{32} in *E.coli* were found for the *rpoH* genes from the γ, but not the α, subgroup, despite considerable divergence in nucleotide sequence. Moreover, a stretch of nine amino acid residues Q(R/K)(K/R)LFFNLR, designated the 'RpoH box', was absolutely conserved among all σ^{32} homologs, but absent in other σ factors; this sequence overlapped with the segment of polypeptide thought to be involved in DnaK/DnaJ chaperone-mediated negative control of synthesis and stability of σ^{32}. In addition, a putative σ^{E} (σ^{24})-specific promoter was found in front of all *rpoH* genes from the γ, but not α, subgroup. These results suggest that the regulatory mechanisms, as well as the function, of the heat shock response known in *E.coli* are very well conserved among the γ subgroup and partially conserved among the α proteobacteria.

INTRODUCTION

Induction of heat shock proteins under heat or other stress conditions represents a ubiquitous cellular response that occurs at the level of transcription (1). In *E.coli* the *rpoH* (*htpR*, *hin*) gene encoding a minor σ factor of 32 kDa (σ^{32}) plays a major role in this homeostatic process (2–4). σ^{32} is not only required for recognizing promoters for transcription of heat shock genes, but serves as the central regulatory factor (5,6). During steady-state growth the level of σ^{32} is kept very low (10–30 molecules/cell) primarily because of limited translation of *rpoH* mRNA and the unusual instability of σ^{32}. Upon a shift to higher temperature the σ^{32} level increases rapidly, though transiently, by both increased synthesis and stabilization. This transient increase in the σ^{32} level essentially accounts for induction of heat shock proteins under stress conditions (5,6).

Extensive analyses of heat-induced synthesis of a σ^{32}–β-galactosidase fusion protein from a *rpoH–lacZ* gene fusion revealed that two proximal coding regions (A and B) of *rpoH* mRNA are involved in translational repression during steady-state growth and marked induction upon temperature upshift, primarily by modulating the stability of the mRNA secondary structure (7–9). Whereas region A (nt 6–20) represents a translational enhancer ('downstream box'; 10) which is complementary to 16S rRNA (nt 1469–1483), region B (nt 110–210) contains a sequence that forms an internal mRNA secondary structure involving the initiation codon and downstream box and thus serves to modulate efficiency of translation (8,9). Furthermore, a segment of σ^{32} polypeptide (region C; around residues 122–144), which corresponds to further downstream of *rpoH*, was found to be important for DnaK/DnaJ chaperone-mediated negative feedback control of the heat shock response by modulating synthesis and stability of the fusion protein (11,12).

On the other hand, no evidence for involvement of σ^{32}-like factors in heat shock gene expression has been found with gram positive bacteria, such as *Bacillus subtilis* (13) and *Clostridium acetobutylicum* (14). Instead, a highly conserved inverted repeat sequence was detected in front of some of the major heat shock genes (*dnaK* and *groE*), which appeared to play a negative regulatory role (13–15). These interesting developments, as well as the unique regulatory features associated with different segments of *rpoH* mRNA and σ^{32} in *E.coli*, prompted us to examine *rpoH* homologs from other bacteria at various phylogenetic distances.

The only sequence of a *rpoH* homolog known when this work was initiated was that of *Citrobacter freundii*, closely related to *E.coli* (16). The *rpoH* gene of more distantly related *Pseudomonas aeruginosa* was recently isolated and characterized (17,18). We have cloned and sequenced the genes from several other bacteria, including some that are even more distantly related. The

* To whom correspondence should be addressed

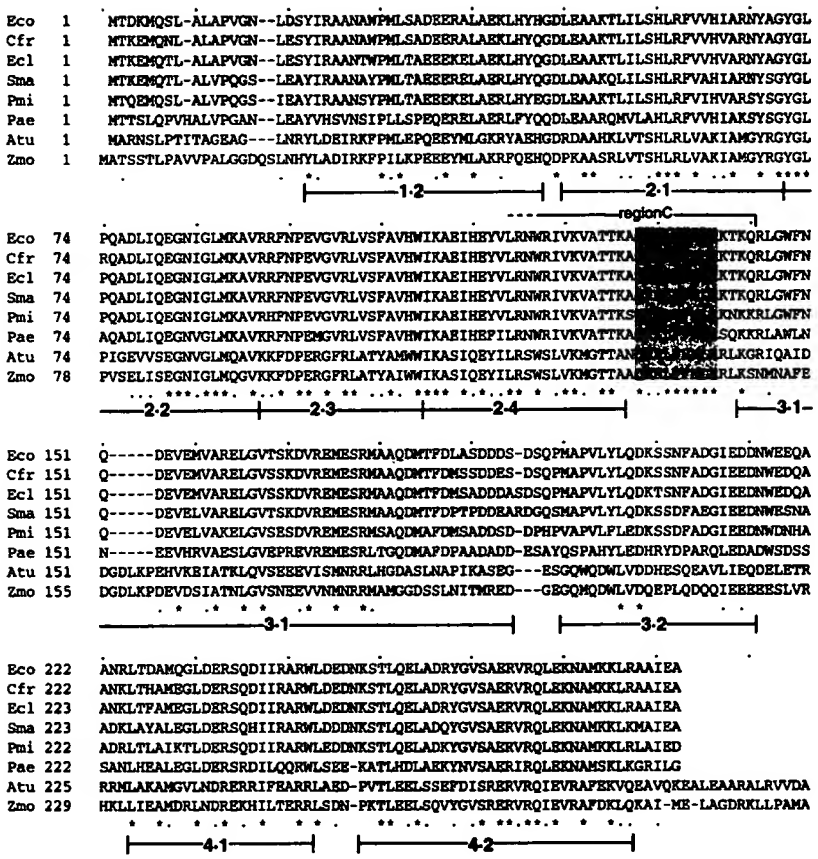

Figure 1. Alignment of deduced amino acid sequences among RpoH homologs. Multiple alignment was carried out with CLUSTAL W (23) using sequences of eight RpoH homologs from *E.coli* (Eco; A94012), *C.freundii* (Cfr; S04697), *E.cloaceae* (Ecl), *S.marcescens* (Sma), *P.mirabilis* (Pmi), *P.aeruginosa* (Pae; U09560), *A.tumefaciens* (Atu) and *Z.mobilis* (Zmo). Numbers below the sequences show generally conserved regions for σ factors according to Lonetto *et al.* (32). Asterisks and dots indicate completely or partially conserved residues, respectively. The shaded area refers to the 'RpoH box' (see text).

selection employed was based on the ability of *rpoH* homologs to functionally complement temperature-sensitive growth of the *E.coli* Δ*rpoH* mutant lacking σ^{32}, which can grow only at or below 20°C (19). Compilation of the resulting sequence data permitted analysis of eight RpoH homologs (σ^{32}-like proteins) altogether from diverse gram negative bacteria.

MATERIALS AND METHODS

Bacteria, phage, plasmids and DNAs

Escherichia coli strain KY1608 (MC4100 Δ*rpoH30*::*kan zhf50*::Tn*10*), a non-lysogenic version of KY1612 described previously (19), was used as the host for initial screening of *rpoH* homologs. Phage λpF13-(*groEp–lacZ*) carrying *lacZ* under control of the *groE* heat shock promoter has been previously described (20). Other phage and plasmid vectors were obtained from commercial sources. The bacteria used as DNA sources were: *Serratia marcescens* ATCC264, *Proteus mirabilis* PM-1, *Enterobacter cloacae* (identified by Dr Akira Yokota, Fermentation Institute, Osaka), *Agrobacterium tumefaciens* IAM12544 (a gift of Drs Kan Tanaka and Hideo Takahashi, University of Tokyo) and *Zymomonas mobilis* CP4 (a gift of Dr L.O.Ingram, University of Florida).

Cloning and sequencing of genes

Cells of strain KY1608 were transfected with a DNA library (>10^4 clones) from each donor bacteria, using charomid 9-36 or λgt11 as a cloning vector. Ampicillin-resistant colonies were selected at 30°C, the transformants obtained were lysogenized with λpF13-(*groEp–lacZ*) and those that exhibited strong red (or blue) color on McConkey lactose agar (or L agar containing X-gal) were further analyzed. The recombinant DNAs containing a putative *rpoH* homolog were subcloned to 2–4 kb fragments by digesting with restriction endonucleases or exonucleases, followed by insertion into plasmid pUC118. Stepwise deletion derivatives of each clone were constructed and the nucleotide sequence was determined with a DNA sequencer (Applied Biosystems). A search for sequence homology was carried out against nucleotide and peptide sequence databases using a BLAST e-mail server (21,22). An open reading frame with a sequence closely related to *E.coli rpoH* was thus detected.

Multiple sequence alignment and phylogenetic analysis

Multiple alignment of sequences was carried out with CLUSTAL W (23) or ICOT Free Software (24), followed by minimum manual modifications. Phylogenetic analysis was conducted with computer programs PROTDIST, NEIGHBOR, PROTPARS, SEQBOOT and CONSENSE, all in the PHYLIP package, version 3.5c (25). Phylogenetic trees were visualized with DRAWGRAM, provided in the same package.

Prediction of RNA secondary structure

Potential mRNA secondary structures for the 5'-portion of the *rpoH* coding region were predicted using MFOLD (26) and the resulting structures visualized with LOOPVIEWER (27).

Accession number for nucleotide sequences

The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL and NCBI nucleotide sequence databases with the following accession numbers. *E.cloacae*, D50829; *S.marcescens*, D50831; *P.mirabilis*, D50830; *A.tumefaciens*, D50828; *Z.mobilis*, D50832.

RESULTS

Cloning and sequencing of *rpoH* homologs

Infection of the *E.coli* mutant lacking σ^{32} (KY1608) with a charomid (or λgt11) library of donor bacterial DNA followed by lysogenization and screening for increased *lacZ* expression, as described in Materials and Methods, led us to identify a *rpoH* gene homolog from each of the five gram negative bacteria belonging to the γ (*E.cloacae*, *S.marcescens* and *P.mirabilis*) or α (*A.tumefaciens* and *Z.mobilis*) subgroups of the proteobacteria. The putative *rpoH* homologs thus obtained contained sequences that are closely related to that of *E.coli rpoH* for the entire coding region (see below). As expected from the selection employed, all the *rpoH* homologs supported growth of KY1608 at 30°C or higher temperatures and permitted heat-induced synthesis of major heat shock proteins. Furthermore, some of them (those from the γ, but not the α, subgroup) caused synthesis of RpoH proteins that can cross-react with antiserum against *E.coli* σ^{32}. Details of the cloning and expression studies will be presented elsewhere (K.Nakahigashi *et al.*, unpublished). The isolation and characterization of the *rpoH* homolog from *P.aeruginosa* was recently reported (17,18).

Overall structural similarity of RpoH homologs

The amino acid sequences predicted from nucleotide sequences of eight *rpoH* homologs, including the five new entries, were aligned to analyze their structures (Fig. 1). The RpoH homologs from the γ subgroup (the upper six genes) were found to be closely related: they showed an almost identical number of amino acid residues (284 or 285) and 60% or higher sequence homology (identity), containing only a few deletions (or additions) of single residues (Fig. 1 and Table 1). This overall similarity in amino acid sequence, despite appreciable differences in nucleotide composition (the GC content varying between 43.1 and 62.7%), was striking. In contrast, the two RpoH homologs from the α subgroup (the lower two genes) were significantly larger (300–302 residues) and contained more extensive deletions or additions, as well as substitutions, thus exhibiting <40% sequence identity with that of *E.coli*. These results are consistent with the known phylogenetic distance between the α and γ proteobacteria (29).

Amino acid sequence comparisons of bacterial σ factors previously permitted identification of four major regions of homology, 1–4, that were further divided into subregions (see 30–32). We thus calculated similarity scores along the entire sequence of the RpoH homologs (data not shown). As might have been expected, regions 2.1–2.4 and 4.2 exhibited particularly high conservation. However, an additional region of similarity appeared to exist between regions 2.4 and 3.1.

A segment(s) uniquely conserved among RpoH homologs

We then compared the above sequence data with those of other σ factors, such as RpoD homologs (primary σ factors, σ^{70} in gram negative bacteria). Inspection of the data suggested that a segment highly conserved among RpoH homologs, but not among the other σ factors, coincides in part with the non-conserved region flanked by regions 2.4 and 3.1. To further substantiate this possibility we compared sequence similarity among RpoH, RpoD and both RpoH and RpoD combined for the entire region spanning from 2.1 to 4.2, using the pair of genes from three representative bacteria, *E.coli, P.aeruginosa and A.tumefaciens* (Fig. 2).

Table 1. Overall structure comparison of the RpoH homologs

	rpoH gene No. of nucleotides	Percent GC	Percent identity (vs *E.coli*)	σ^{32} homolog No. of amino acids	Mol. wt	Percent identity (vs. *E.coli*)
E.coli	852	54.3	100	284	32.5	100
C.freundii	852	52.9	89.9	284	32.6	94.4
E.cloacae	855	54.6	86.5	285	32.7	92.3
S.marcescens	855	57.8	80.7	285	32.7	84.9
P.mirabilis	852	43.1	72.8	284	32.6	80.3
P.aeruginosa	852	62.7	66.4	284	32.6	60.7
A.tumefaciens	900	60.4	45.4	300	34.4	35.6
Z.mobilis	906	49.4	46.4	302	34.5	38.6

All the numbers and values refer to the coding sequences. Percent identity was based on a comparison with *E.coli rpoH* or σ^{32}. Pairwise alignments of either nucleotide or amino acid sequences were carried out using the ALIGN 0 program, version 17 (28).

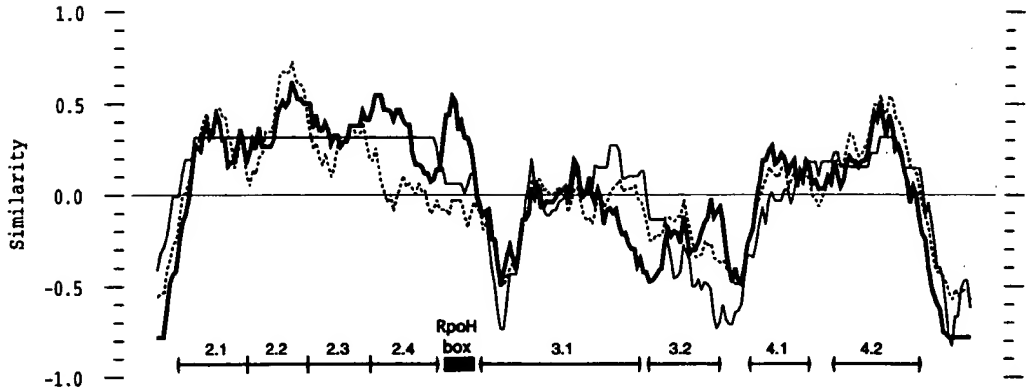

Figure 2. Sequence similarity among the RpoH and RpoD families of σ factors. The sequence data of Figure 1 (RpoH homologs) and that of the RpoD family (*E.coli* RpoD, A00699; *E.coli* RpoS/KatF, S14901; *P.aeruginosa* RpoD, S15900; *A.tumefaciens* SigA, A36913; *M.xanthus* SigA, M32347; *B.subtilis* RpoD, A22626; *P.aeruginosa* RpoS, D26134; *Lactococcus lactis* RpoD, JC1397) were aligned with ICOT Free Software (24). Three representative bacteria in which both RpoD and RpoH sequences are known (*E.coli, P.aeruginosa* and *A.tumefaciens*) were then used to calculate similarity scores using PLOTSIMILARITY (in WISCONSIN PACKAGE version 8, Genetics Computer Group Inc.). Thick and thin lines show similarity within the RpoH or RpoD families, respectively, whereas the dotted line shows similarity with both RpoH and RpoD sequences combined. The scores have been normalized to the average value (0.0) for each plot (0.8, 1.2 and 0.7 for the above three lines, respectively). A plateau level for the conserved 2.1–2.4 region of RpoD represents the maximum value obtained (perfect match). A narrow region between 2.4 and the RpoH box with apparently low similarity for RpoH is due to an artifactual gap introduced during co-alignment of the RpoH and RpoD sequences.

In the case of RpoD the very highly conserved region 2.1–2.4 was immediately followed by a region of much lower similarity. With RpoH, however, the conserved region 2.4 was followed by another region of high similarity, comparable with that of the preceding region. In particular, a stretch of nine amino acids Q(R/K)(K/R)LFFNLR (residues 132–140 for *E.coli*) was the most highly conserved in the entire sequence and was tentatively designated the 'RpoH box'. In contrast, pairwise comparisons of the nucleotide sequence for the same segment revealed appreciable variation between different species (up to eight changes in 27 nucleotides; not shown). Furthermore, similarities for region 2.4 and the RpoH box found with the combined RpoD and RpoH sequences were much lower than those obtained with RpoH (or RpoD) alone, indicating that these two regions represent unique and characteristic features of RpoH homologs. Comparisons between RpoH and other σ factors also gave results that are consistent with the notion that the RpoH box represents a sequence unique to this family of σ factors (not shown).

Most significantly, the RpoH box overlapped with region C (around residues 122–144), which was previously shown to be involved in DnaK/DnaJ-mediated translational shut-off of σ^{32} synthesis during the adaptation phase and in the characteristic instability of σ^{32}, on the basis of observations with deletion and frame-shift derivatives of a *rpoH–lacZ* gene fusion (11). Thus region C, which contains at least part of region 2.4 and perhaps the entire RpoH box, is very well conserved among all the RpoH homologs examined (see Fig. 1).

Conserved regulatory elements for translational control

A sequence similar to the downstream box, which is complementary to 16S rRNA, was found immediately following the initiation codon of *rpoH* mRNAs for all the γ proteobacteria analyzed (Fig. 3). The complementarity was as high as that previously found in *E.coli* (65–80% matching), strongly suggesting that these sequences play an active role in enhancing translation of the

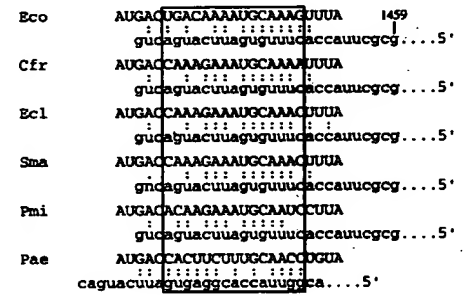

Figure 3. The 5'-portions (nt 1–24) of coding sequence containing the putative 'downstream box' of each *rpoH* homolog (upper sequence) were shown to be complementary to part of 16S rRNA ('anti-downstream box, spanning nt 1469–1483 in *E.coli*, close to the 3'-end; lower sequence) of the respective bacteria. The sequence data used, besides those listed in the legend to Figure 1, were: Eco *rpoH*, X04398; Eco 16S rRNA (see 10); Cfr *rpoH*, X14960; Cfr 16S rRNA, M59291; Sma 16S rRNA, M59160; Pvu 16S rRNA, X07652; Pae 16S rRNA, M34133. The anti-downstream box of *P.aeruginosa* is shifted by eight bases toward its 5'-end, as compared with that of the others. Because the 16S rRNA sequences for *E.cloacae* and *P.mirabilis* are not known, those from close relatives (*E.coli* and *Proteus vulgaris*, respectively) were used.

respective *rpoH* mRNA in these bacteria, in addition to the similar role played by the Shine–Dalgarno sequence. It should be noted that the anti-downstream box that matches with *P.aeruginosa rpoH* was slightly shifted from those of the rest of the bacteria examined. This probably explains the previous failure to detect such a sequence (18). No similar sequence was found with the *A.tumefaciens* gene, and the lack of 16S rRNA sequence data prevented analysis of the *Z.mobilis* gene.

We then examined potential mRNA secondary structures for the 5'-proximal portion (nt −20 to 210) of the *rpoH* coding sequence, which could be involved in thermal regulation of its translation. The predicted mRNA secondary structures for all the

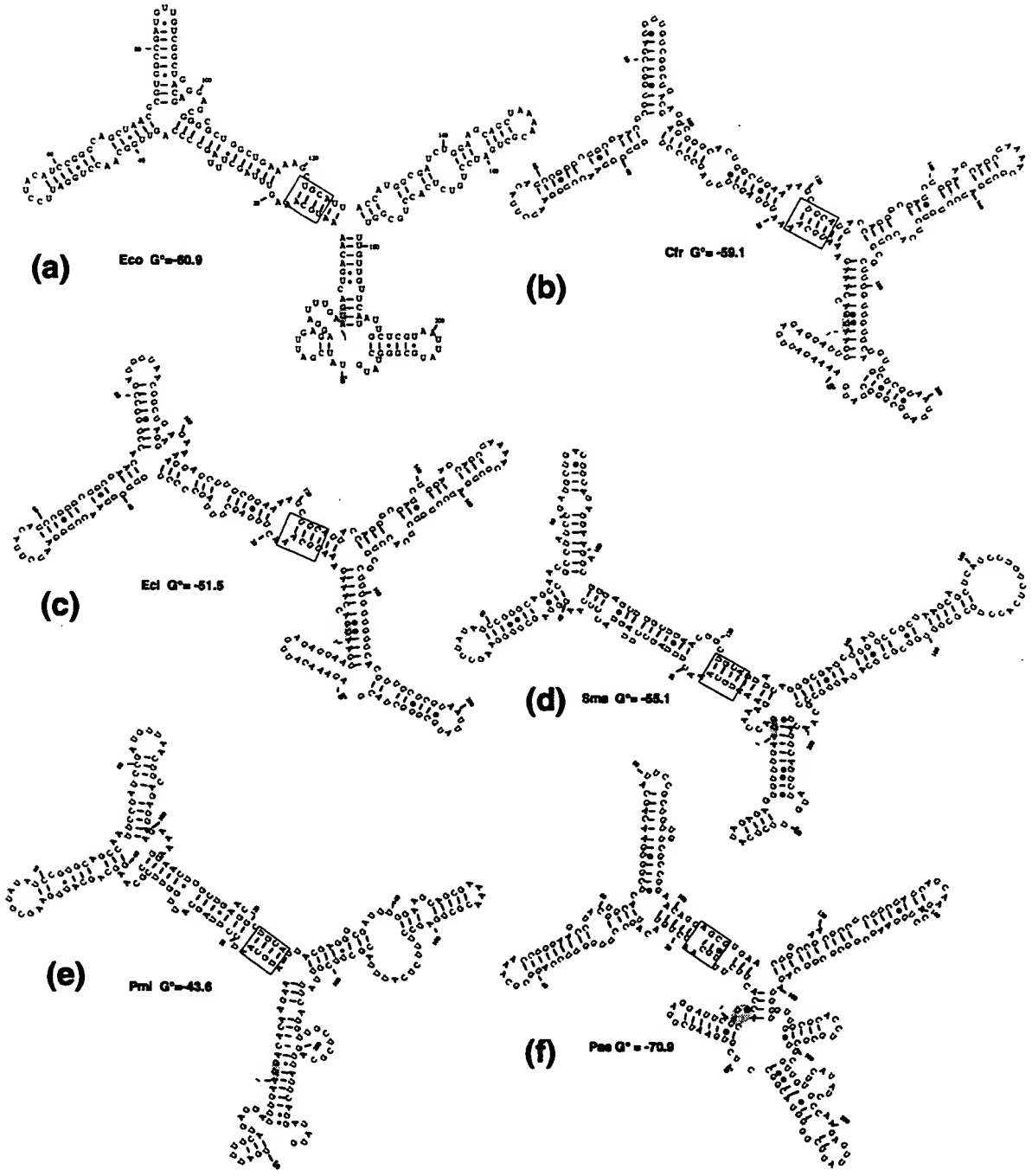

Figure 4. Possible secondary structures for the 5'-portion of mRNA (nt –20 to 210) of *rpoH* homologs from γ proteobacteria. The structures shown were among the most stable predicted by MULFOLD (26) and visualized by LOOPVIEWER (27) with one exception, *P.mirabilis*. The minimum energies (ΔG, kcal/mol) were calculated to be –60.9 (Eco), –59.4 (Cfr), –51.5 (Ecl), –55.1 (Sma), –53.5 (Pmi) and –70.9 (Pae), which should be compared with the values indicated for each of the structures shown. The initiation codon is shaded and the conserved base pairings are boxed (see text).

γ proteobacteria were surprisingly similar to that reported for *E.coli*; the structures presented in Figure 4 were found as one of the most stable structures for each of the *rpoH* homologs examined (except for *P.mirabilis*); the three major stems thought to play unique regulatory roles (9) were well conserved. The similarity extended further, to, apparently, the two most critical adjacent

```
Eco  ttGAACTTgtggataaaatcactgTCTGAtaaaa   76   GAGAGGAttgaATG
Cfr        Tgtggataaaatcactg TCTGAtaaaa   92   GAGAGGAttgaATG
Ecl  tgGAACTTgtggataaaatcactg TCTGAtaaaa   89   GAGAGGAat-gaATG
Sma  gtGAACTTttagggcggagtacgg TCAAAaattg  121   GAGAGGGttgaATG
Pmi  atGAACCTttagtctttatgctgg TCAATaatgt  115   ATGAGGAttgaATG
Pae  agGAACTTatacacccgcttgcag TCAGAtatcc   30   CGGAGGAttcgcATG
          -35                  -10           SD
```

Figure 5. Predicted σ^E promoters found in the upstream region of *rpoH* homologs are shown along with that known in *E.coli* (see 6) and that predicted for *P.aeruginosa* (17,18). SD, Shine–Dalgarno sequence. The numbers refer to nucleotides that should be inserted between the pomoter region and SD sequence.

G-C pairings involved in thermal regulation of *E.coli rpoH* (nt 15::124 and 16::123): they were perfectly conserved among all the mRNA structures shown (shaded areas in Fig. 4). These results, along with the recent report on *P.aeruginosa rpoH* (18), provide strong evidence in support of the notion that mRNA secondary structure plays an important role in the control of *rpoH* translation in *E.coli* and, by inference, in other members of the γ proteobacteria examined.

Comparison of flanking nucleotide sequences

Among the known promoters of *E.coli rpoH*, three are transcribed by σ^{70} RNA polymerase ($E\sigma^{70}$) and one by σ^E RNA polymerase ($E\sigma^E$) (12). A search for putative promoters with the *rpoH* homologs led us to find a putative σ^E promoter in front of each of the *rpoH*s of all the γ proteobacteria (Fig. 5). In contrast, we failed to identify any promoter(s) similar to the 'σ^{70}-consensus', although at least one such promoter is probably active. In *P.aeruginosa* two major *rpoH* transcripts were reported to be produced when the gene was expressed in *E.coli* upon shift to 50°C, though the start site for a putative σ^E-dependent transcript did not quite agree with the predicted location of the promoter (17,18). On the basis of sequence data alone we failed to detect either σ^{70}-like or σ^E-like promoters upstream of the *rpoH* homologs from the α subgroup (*A.tumefaciens* and *Z.mobilis*). Also, none of the *rpoH* genes examined contained promoter sequences similar to the σ^{32}-consensus. On the other hand, a typical transcription terminator sequence, consisting of a hairpin structure followed by a T(U) cluster, was found shortly after the termination codon on mRNAs of all the *rpoH* homologs analyzed (not shown).

Phylogenetic relationships among RpoHs and other σ factors

Finally, we analyzed phylogenetic relationships among RpoH homologs, as well as other σ factors of gram negative bacteria, including RpoD (σ^{70}), RpoS (σ^S), RpoE (σ^E) and RpoF (σ^{28}). RpoN (σ^{54}) was excluded from this analysis, because it was known to be virtually unrelated to any of the other σ factors (see 32). The generally conserved region 2.1–2.4 was compared using two independent methods, which yielded essentially identical results (Fig. 6). It seems evident that the RpoH homologs form a distinct cluster, separate from the rest of the σ factors, suggesting that the RpoH homologs originated from a common ancestor during evolution. The *sigB* and *sigC* gene products of *Myxococcus xanthus* were apparently most closely related phylogenetically to the RpoH homologs among the σ factors whose sequences are currently available in the databases.

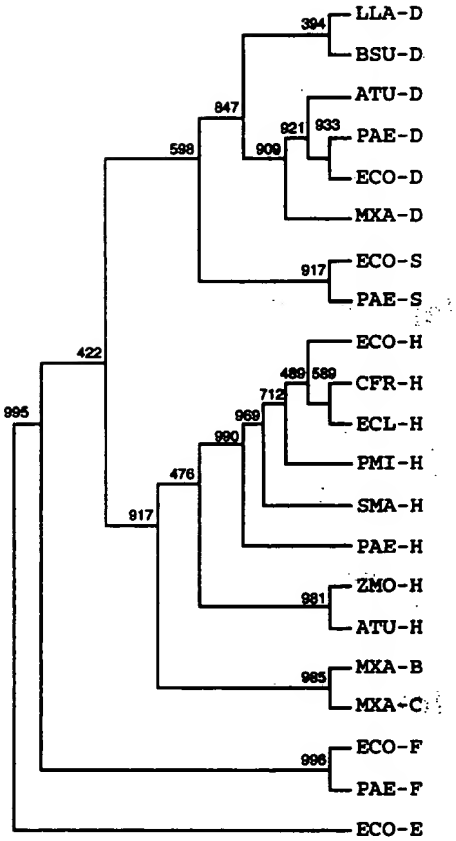

Figure 6. Consensus phylogenetic trees of σ factors from gram negative bacteria were analyzed by the neighbor joining and parsimony methods. The entire sequences of region 2 were aligned with ICOT Free Software and used to obtain a consensus tree after 1000 bootstrap replications. Numbers given near the forks indicate calculated probabilities (per thousand) that the given species shown to the right should be grouped together, by means of a 1000 bootstrap analysis. The sequence data used included those used in Figures 1 and 2 plus *M.xanthus* SigB (MXA-B, X55500), *M.xanthus* SigC (MXA-C, L12992), *E.coli* FliA (ECO-F, P31804), *P.aeruginosa* FliA (PAE-F, S20544) and *E.coli* RpoE (ECO-E, P34086). Only the results from the neighbor joining analysis are shown; essentially identical results were obtained in the parsimony analysis.

DISCUSSION

The present analysis of RpoH homologs from the α and γ subgroups of the proteobacteria revealed that they form a distinct cluster in the phylogenetic tree among the σ factors of gram negative bacteria (Fig. 6). This is consistent with the observation that these homologs can closely mimic the function of σ^{32} when expressed in *E.coli* (K.Nakahigashi *et al.*, unpublished) and suggests that the regulatory, as well as catalytic, functions of σ^{32} are well conserved among these bacteria. The only other σ factors closely related to RpoH were SigB (MxaB) and SigC (MxaC) of *M.xanthus*, which belongs to the δ subgroup of the proteobacteria (see Fig. 6; 31). However, when the genes for these σ factors were prepared by PCR and introduced into the *E coli* Δ*rpoH* mutant they failed to complement temperature-sensitive growth of the mutant, even at 30°C (K.Nakahigashi, unpublished).

The segments of σ^{32} particularly well conserved among these homologs, but distinct from those of other σ factors, overlapped with region 2.4 and an adjacent region which contained nine residues of the 'RpoH box' (Fig. 1). The high conservation of region 2.4 specifically among the RpoH homologs was not unexpected, because this region is known to be primarily responsible for recognition of −10 promoter sequences (32). In agreement with this, all the RpoH homologs tested were capable of promoting synthesis of at least some of the major heat shock proteins when expressed in the *E.coli* Δ*rpoH* mutant (K.Nakahigashi *et al.*, unpublished).

Although the RpoH box located outside region 2.4 might also reflect a unique promoter specificity of this σ factor, it seems most likely that this region is involved in regulation characteristic of the RpoH homologs, because it was found to overlap with region C of *E.coli* σ^{32}, supposed to be critical for DnaK/DnaJ-mediated negative control of its synthesis and degradation (11). It has been shown that strains carrying the *rpoH–lacZ* gene fusion with a deletion or frame-shift mutation affecting region C failed to shut off synthesis of fusion protein during the adaptation phase of the heat shock response (11,12). Moreover, the fusion proteins produced from these mutant derivatives are very stable, unlike the parental fusion protein, which is as unstable as authentic σ^{32} (11). Recently it was found that peptides of 13 amino acids that contain the RpoH box actually bind to DnaK with the highest affinities among a set of overlapping peptides spanning the entire σ^{32} polypeptide (J.McCarty and B.Bukau, personal communication)

We propose that the RpoH box, perhaps with its flanking sequence(s), is specifically required for chaperone-mediated negative control of the synthesis and/or degradation of RpoH proteins in a variety of gram negative bacteria. This would imply that the RpoH protein itself is involved in regulation of the heat shock response in all these bacteria, as had been documented in *E.coli* (6,12). This is interesting, but perhaps not surprising, because feedback control of the heat shock response is likely to be universal, involving transcription factors such as σ^{32} and eukaryotic heat shock factors (HSF) on the one hand and chaperones such as DnaK/DnaJ and HSP70 on the other, not only in prokaryotes but also in many eukaryotic organisms (1,12).

Besides the RpoH box and region C discussed above, the 5′-segment of mRNA immediately downstream of the initiation codon (downstream box; Fig. 3) and the mRNA secondary structure with its major stems (Fig. 4) is highly conserved among most (if not all) members of the γ proteobacteria, despite the wide variation in nucleotide composition and sequence among some of these homologs. These results strongly suggest that translational repression of RpoH proteins, as mediated by mRNA secondary structure during steady-state growth, and the rapid and transient activation following heat shock is a well-conserved mode of regulation among these bacteria. It is therefore likely that the regulatory mechanisms of the heat shock response in these bacteria are quite similar to what has been found in *E.coli*. A similar analysis with a putative *rpoH* homolog of *Haemophilus influenzae* Rd (33), an additional member of the γ subgroup, agreed well with this expectation (K.Nakahigashi *et al.*, unpublished).

In the α proteobacteria, including *A.tumefaciens* and *Z.mobilis*, transcription of a number of heat- or ethanol-inducible genes appears to be initiated from promoters similar to the 'σ^{32}-consensus' found in *E.coli*. This includes *virG* of *A.tumefaciens* (34), *dnaK* of *Caulobacter crescentus* (35) and *adhB* of *Z.mobilis* (36). Besides, the *E.coli dnaK* heat shock gene was shown to be transcribed from known heat shock promoter(s) when expressed in *A.tumefaciens*, suggesting the involvement of a σ^{32}-like factor in this bacterium (37). On the other hand, transcription of the *groE* (*groES–groEL*) operon in *A.tumefaciens* was reported to be initiated from a σ^{70}-like promoter before or after heat shock (38). More direct experiments *in vivo* and *in vitro* would be required to resolve the apparent discrepancy between this and the other experiments cited above.

It should also be noted that a well-conserved inverted repeat sequence similar to that found in gram positive bacteria is present in front of the *groE* operon in both of the α proteobacteria examined, *A.tumefaciens* and *Z.mobilis* (38,39). Since such inverted repeats have so far been detected only in the *groE* and *dnaK* operons in gram positive or negative bacteria, their roles in heat shock regulation may be restricted to genes encoding some of the major chaperones (15). The functional interplay of the positive (σ^{32}) and negative elements (inverted repeats) in heat shock regulation of α proteobacteria would be an interesting subject for further study. Finally, extension of the present work to include additional members of the proteobacteria, as well as other distantly related bacteria, may provide useful information on the phylogenetic distribution of σ^{32} and other regulatory factors involved in the heat shock response in prokaryotes.

ACKNOWLEDGEMENTS

We are grateful to Drs Akira Yokota, Hideo Takahashi, Kan Tanaka and L. O. Ingram for bacterial strains and DNAs, to Drs Takashi Miyata and Hideo Matsuda for advice on computer sequence analysis and to Dr Masato Ishikawa for use of some computer programs. Technical assistance by Ms Masako Nakayama and Mayumi Ueda is also appreciated. This work was supported in part by grants from the Japan Health Sciences Foundation, Tokyo.

REFERENCES

1 Morimoto,R., Tissires,A. and Georgopoulos,C. (1994) *The Biology of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor Laboratory Press, Plainview, NY, p. 610.
2 Grossman,A.D., Erickson,J.W. and Gross,C.A. (1984) *Cell*, **38**, 383–390.
3 Landick,R., Vaughn,V., Lau,E.L., VanBogelen,R.A., Erickson,J.W. and Neidhardt,F.C. (1984) *Cell*, **38**, 175–182.
4 Yura,T., Tobe,T., Ito,K. and Osawa,T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6803–6807.
5 Straus,D., Walter,W. and Gross,C.A. (1987) *Nature*, **329**, 349–351.
6 Gross,C.A., Straus,D.B., Erickson,J.W. and Yura,T. (1990) In Morimoto,R., Tissires,A. and Georgopoulos,C. (eds), *Stress Proteins in Biology and Medicine*. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. 167–189.
7 Kamath-Loeb,A.S. and Gross,C.A. (1991) *J. Bacteriol.*, **173**, 3904–3906.
8 Nagai,H., Yuzawa,H. and Yura,T. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 10515–10519.
9 Yuzawa,H., Nagai,H., Mori,H. and Yura,T. (1993) *Nucleic Acids Res.*, **21**, 5449–5455.
10 Sprengart,M.L., Fatscher,H.P. and Fuchs,E. (1990) *Nucleic Acids Res.*, **18**, 1719–1723.
11 Nagai,H., Yuzawa,H., Kanemori,M. and Yura,T. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 10280–10284.
12 Yura,T., Nagai,H. and Mori,H. (1993) *Annu. Rev. Microbiol.*, **47**, 321–350.
13 Wetzstein,M., Vslker,U., Dedio,J., Lobau,S., Zuber,U., Schiesswohl,M., Herget,C., Hecker,M. and Schumann,W. (1992) *J. Bacteriol.*, **174**, 3300–3310.
14 Narbenhaus,F., Giebeler,K. and Bahl,H. (1992) *J. Bacteriol.*, **174**, 3282–3289.
15 Zuber,U. and Schumann,W. (1994) *J. Bacteriol.*, **176**, 1359–1363.
16 Daggett Garvin,L. and Hardies,S.C. (1989) *Nucleic Acids Res.*, **17**, 4889.

17 Naczynski,Z.M., Mueller,C. and Kropinski,A.M. (1995) *Can. J. Microbiol.*, **41**, 75–87.
18 Benvenisti,L., Koby,S., Rutman,A., Giladi,H., Yura,T. and Oppenheim,A.B. (1995) *Gene*, **155**, 73–76.
19 Zhou,Y.-N., Kusukawa,N., Erickson,J.W., Gross,C.A. and Yura,T. (1988) *J. Bacteriol.*, **170**, 3640–3649.
20 Yano,R., Imai,M. and Yura,T. (1987) *Mol. Gen. Genet.*, **207**, 24–28.
21 Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990). *J. Mol. Biol.*, **215**, 403–410.
22 Gish,W. and States,D.J. (1993) *Nature Genet.*, **3**, 266–272.
23 Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res.*, **22**, 4673–4680.
24 Ishikawa,M., Toya,T. and Totoki,Y. (1994) *Proceedings of the International Symposium on Fifth Generation Computer Systems '94*, pp. 129–138.
25 Felsenstein,J. (1991) PHYLIP manual, version 3.5, University of Washington, Seattle, WA.
26 Jaeger,J.A., Turner,D.H. and Zuker,M. (1990) *Methods Enzymol.*, **183**, 281–306.
27 Gilbert,D.G. (1990) Internet ftp://iubio.bio.indiana.edu..
28 Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2444–2448.
29 Woese,C.R. (1987) *Microbiol. Rev.*, **51**, 221–271.
30 Helmann,J.D. and Chamberlin,M. (1988) *Annu. Rev. Biochem.*, **57**, 839–872.
31 Lonetto,M., Gribskov,M. and Gross,C.A. (1992) *J. Bacteriol.*, **174**, 3843–3849.
32 Gross,C.A., Lonetto,M. and Losick,R. (1992) In McKnight,S.L. and Yamamoto,K.R. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. 129–176.
33 Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J.-F., Dougherty,B.A., Merrick,J. M. *et al.* (1995) *Science*, **269**, 496–512.
34 Mantis,N.J. and Winans,S.C. (1992) *J. Bacteriol.* **174**, 1189–1196.
35 Gomes,S.L., Gober,J.W. and Shapiro,L. (1990) *J. Bacteriol.*, **172**, 3051–3059.
36 An,H., Scopes,R.K., Rodriguez,M., Keshav,K.F. and Ingram,L.O. (1991) *J. Bacteriol.*, **173**, 5975–5982.
37 Mantis,N.J. and Winans,S.C. (1992) *J. Bacteriol.*, **174**, 991–997.
38 Segal,G. and Ron,E.Z. (1993) *J. Bacteriol.*, **175**, 3083–3088.
39 Barbosa,M. de F.S., Yomano,L.P. and Ingram,L.O. (1994) *Gene*, **128**, 51–57.